

# Rapid mixing of subset Glauber dynamics on graphs of bounded tree-width

Magnus Bordewich      Ross J. Kang

January 12, 2013

## Abstract

Motivated by the ‘subgraphs world’ view of the ferromagnetic Ising model, we develop a general approach to studying mixing times of Glauber dynamics based on subset expansion expressions for a class of graph polynomials. With a canonical paths argument, we demonstrate that the chains defined within this framework mix rapidly upon graphs of bounded tree-width. This extends known results on rapid mixing for the Tutte polynomial, the adjacency-rank ( $R_2$ -)polynomial and the interlace polynomial.

Keywords: Markov chain Monte Carlo, subset expansion, graph polynomials, tree-width, canonical paths, Tutte polynomial, interlace polynomial, random cluster model.

## 1 Introduction

We analyse a subset-sampling Markov chain on simple graphs that is derived from certain graph functions — usually, in fact, graph polynomials. We show that this chain mixes rapidly on graphs of constant tree-width.

Throughout the paper, the graph functions  $\mathcal{P}$  we consider are formulated using subset expansion<sup>1</sup>. An *edge subset expansion formula for  $\mathcal{P}$*  is written as follows: for any simple graph  $G = (V, E)$ ,

$$\mathcal{P}(G) = \sum_{S \subseteq E} w((V, S)) \quad (1)$$

for some graph function  $w$ , where  $(V, S)$  denotes the graph with vertex set  $V$  and edge set  $S$ . If the function  $w$  is non-negative, that is,  $w(G) \geq 0$  for all graphs  $G$ , we refer to (1) as an *edge subset weighting for  $\mathcal{P}$*  and to  $w$  as its *weight function*. In fact, we shall need the weight function to be *positive* on all subgraphs — from a statistical physics viewpoint, this results in a so-called ‘soft-core model’.

Before moving on, let us anchor the general formula (1) with an example that is prominent in statistical physics, theoretical computer science, and discrete probability. The *partition function of the random cluster model* can be defined

---

<sup>1</sup>The term ‘subset expansion’ was coined by Gordon and Traldi [28], though it is a special type of ‘states model expansion’ which is commonly used in physics.

for any  $G = (V, E)$  and parameters  $q, \mu$  as

$$Z_{RC}(G; q, \mu) := \sum_{S \subseteq E} q^{\kappa(S)} \mu^{|S|}, \quad (2)$$

where  $\kappa(S)$  is the number of components in  $(V, S)$ . For more on the random cluster model, see an extensive treatise by Grimmett [29]. Notice that, if  $q, \mu \geq 0$ , then  $w((V, S)) := q^{\kappa(S)} \mu^{|S|}$  provides an edge subset weighting for  $Z_{RC}(G; q, \mu)$ . Under a suitable transformation,  $Z_{RC}(G; q, \mu)$  is equivalent to the Tutte polynomial [57], defined for any  $G = (V, E)$  and parameters  $x, y$  as

$$T(G; x, y) := \sum_{S \subseteq E} (x - 1)^{r(E) - r(S)} (y - 1)^{|S| - r(S)}, \quad (3)$$

where  $r(S)$  is the  $\mathbb{F}_2$ -rank of the incidence matrix for  $(V, S)$ . A wealth of combinatorial and structural information can be obtained from evaluations of this function. Indeed, this polynomial has a remarkable universality property, which informally speaking says that it subsumes any graph invariant that can be computed by deletion and contraction of edges [50], cf. [59]. In addition, the Tutte polynomial specialises to several key univariate graph polynomials, including the chromatic polynomial of Birkhoff [7]. It specialises to the Jones polynomial in knot theory [38]. By its connection with the random cluster model, it also generalises the partition functions of the Ising [33] and Potts [51] models<sup>2</sup>. Consult the monograph of Welsh [58] for more on these crucial connections. In addition to  $Z_{RC}(G; q, \mu)$  and  $T(G; x, y)$ , we shall highlight a few other specific polynomials from the literature, but for a broad account of the development of graph polynomials, consult the recent surveys by Makowsky [43] and by Ellis-Monaghan and Merino [19, 20].

It was shown in 1990 by Jaeger, Vertigan and Welsh [34] that, in general, for fixed (rational) values of  $x$  and  $y$ , the evaluation of  $T(G; x, y)$  is #P-hard, except on a few special points and curves in the  $(x, y)$ -plane. As a result, there have been substantial efforts since then to pin down the *approximation complexity* of computing  $T(G; x, y)$ . For large swaths of the  $(x, y)$ -plane, it is now known that the computation of  $T(G; x, y)$  either does not admit a fully polynomial-time randomised approximation scheme (FPRAS) unless  $\text{RP} = \text{NP}$ , or is at least as hard as #BIS (the problem of counting independent sets in bipartite graphs) under approximation-preserving reductions, cf. Goldberg and Jerrum [26]. The sole positive approximation result applicable to general graphs is the breakthrough FPRAS by Jerrum and Sinclair [36, 37] for the partition function of the ferromagnetic Ising model — this corresponds to computation of  $T(G; x, y)$  along the portion of the parabola  $(x - 1)(y - 1) = 2$  with  $y > 1$ . Various approaches have given efficient approximations in some regions of the Tutte plane for specific classes of graphs — cf. e.g. Alon, Frieze and Welsh [2, 3], Karger [39, 40], and Bordewich [13]. To obtain their seminal result, Jerrum and Sinclair used a Markov chain Monte Carlo (MCMC) method, a principal tool in the design of efficient approximation schemes for counting problems. MCMC methods are widespread in computational physics, computational biology, machine learning, and statistics. There have been steady advances in our understanding of such

---

<sup>2</sup>If  $x, y \geq 1$  or  $q, \mu \geq 0$ , then, respectively,  $T(G; x, y)$  or  $Z_{RC}(G; q, \mu)$  generalise the partition functions of the *ferromagnetic* Ising and Potts models.

random processes and in showing how quickly they produce good approximations of useful probability distributions in huge, complex data sets. See the lecture notes of Jerrum [35] or a survey by Randall [52] for an overview of the application of these techniques in theoretical computer science.

We postpone the precise statement of our main result, Theorem 1, as it requires a host of definitions, but here we give a cursory description. In this paper, we are interested in the rate of convergence to stationarity of a natural Markov chain closely associated to a subset weighting of  $\mathcal{P}$  (of form (1)), when some mild restriction is placed upon the weight function  $w$ . That restriction — which we have dubbed  *$\lambda$ -multiplicative* — is described in Subsection 2.1: for now, we remark that some important graph polynomials and partition functions from statistical physics (e.g.  $Z_{RC}(G; q, \mu)$  and  $T(G; x, y)$ ) obey it. The state space of our chain is the set of all edge subsets, upon which we have set up a MCMC method using Glauber dynamics [24]. Each possible transition in the chain is either the addition or deletion of exactly one edge to/from the subset and the transition probabilities are defined according to the weights  $w((V, S))$ , subject to a Metropolis-Hastings filter [30, 46]<sup>3</sup>. Our main finding is that on graphs of bounded tree-width this Markov chain converges to the stationary distribution in time that is polynomial in the number of vertices of the graph.

Our approach is inspired in part by the ‘subgraphs world’ in which Jerrum and Sinclair [36, 37] designed their FPRAS for the partition function of the ferromagnetic Ising model. It is also motivated by recent work of Ge and Štefankovič [22, 23], who introduced the  $R_2$ -polynomial in an attempt to devise a FPRAS for #BIS. Their *adjacency-rank polynomial* is defined for any  $G = (V, E)$  and parameters  $q, \mu$  as

$$R_2(G; q, \mu) := \sum_{S \subseteq E} q^{\text{rk}_2(S)} \mu^{|S|}, \quad (4)$$

where  $\text{rk}_2(S)$  is the  $\mathbb{F}_2$ -rank of the *adjacency* matrix for  $(V, S)$ . Using a combinatorial interpretation of  $\text{rk}_2$  applicable only to bipartite graphs, they showed that the edge subset Glauber dynamics (using the weighting in (4)) mixes rapidly on trees. They conjectured that the chain mixes rapidly on all bipartite graphs, cf. Conjecture 1 in [22]. In addition, Ge and Štefankovič showed that the Markov chain for the (soft-core) random cluster model — i.e. weighted according to (2) — mixes rapidly upon graphs of bounded tree-width. We have extended both of these results under a unified framework. In particular, we show that the  $R_2$ -polynomial fits in our framework without recourse to the combinatorial interpretation for bipartite graphs, and hence that the Markov chain for the  $R_2$ -polynomial mixes rapidly upon all graphs of bounded tree-width. We also remark here that the conjectured rapid mixing of this chain on all bipartite graphs was disproved by Goldberg and Jerrum [25].

The polynomials and Markov chains that we capture in our framework are defined for any graph; however, we obtain rapid mixing results only on graphs of constant tree-width. For brevity, we will not define tree-width here, but merely say that it is an essential concept in structural graph theory and parameterised complexity — see modern surveys on the topic by Bodlaender [12] and Hliněný

---

<sup>3</sup>A Metropolis-Hastings filter is applied in order to ensure that the resulting process is a reversible Markov chain and thus guaranteed to converge to a unique stationary distribution with state probabilities proportional to the weight.

et al. [31]. The restriction of tree-width is commonly used in graph algorithms to reduce the complexity of a computationally difficult problem, usually by way of dynamic programming. For example, it is already known that many of the polynomials covered here can be evaluated efficiently for graphs of bounded tree-width. Independently, Andrzejak [4] and Noble [47] exhibited polynomial-time algorithms to compute the Tutte polynomial of graphs with bounded tree-width. Works of Makowsky and Mariño [44] and Noble [48] have significantly generalised this, in the former case, to a wide array of polynomials under the framework of monadic second order logic (MSOL), and, in the latter case, to the so-called  $U$ -polynomial [49], a polynomial that includes not only the Tutte polynomial but also a powerful type of knot invariant as a special case.

Even though many of the polynomials we refer to can be computed exactly in polynomial time for graphs of bounded tree-width, it remains of interest to show that the associated Glauber dynamics is rapidly mixing. One hope is that for some polynomials the chain mixes rapidly for a wider class of graphs. There have been significant and concerted endeavours by researchers spanning physics, computer science and probability to determine the mixing properties of Glauber dynamics on many related Markov chains. Spin systems have been of particular interest; indeed, the main thrust of the work of Jerrum and Sinclair [36, 37] was to tackle the partition function for the ‘spins world’ of the ferromagnetic Ising model (using a translation to the rapidly mixing ‘subgraphs world’). For more on the connections among the ‘spins world’, the ‘subgraphs world’ and the ‘random cluster world’, see the recent work of Huber [32]. We note that many recent projects on spin systems have been restricted to trees or tree-like graphs, cf. e.g. [6, 16, 18, 27, 45, 55].

Our primary focus in this paper is to establish results for polynomials defined according to *edge* subset expansion, but we can also extend our methodology to polynomials defined according to *vertex* subset expansion, which may be viewed as the ‘induced subgraphs world’. To our knowledge, this form of Markov chain has not been greatly examined, but it handles one important graph polynomial that was recently introduced by Arratia, Bollobás and Sorkin [5]: the *bivariate interlace polynomial*, defined for any graph  $G = (V, E)$  and parameters  $x, y$  as

$$q(G; x, y) := \sum_{S \subseteq V} (x - 1)^{\text{rk}_2(S)} (y - 1)^{|V| - \text{rk}_2(S)}, \quad (5)$$

where  $\text{rk}_2(S)$  is the  $\mathbb{F}_2$ -rank of the adjacency matrix for  $G[S]$ . This polynomial specialises to the independence polynomial and is intimately related to Martin polynomials [1]. Just as for the Tutte polynomial, computation of the bivariate interlace polynomial is  $\#P$ -hard in almost the entire plane, as was shown by Bläser and Hoffmann [9]. The multivariate interlace polynomial, a generalisation of the interlace polynomial, can be evaluated efficiently for graphs of bounded tree-width, cf. Courcelle [15] and Bläser and Hoffmann [10, 8]. Subject to a condition on the vertex subset weightings, which we have called *vertex  $\lambda$ -multiplicativity*, we can establish rapid mixing for vertex subset Glauber dynamics on graphs of constant tree-width.

For all of our results, we need that the weight function is strictly positive for all (induced) subgraphs. Many of the classical enumeration polynomials such as the matching, independence, clique and chromatic polynomials are captured by the general polynomials that we mention as examples throughout this work.

However, these are ‘hard-core models’, in which some (induced) subgraphs have a zero weighting, and hence are not included in our approach. Many of these are evaluations that fall at the boundary of the regions that we can handle. For example, the Tutte polynomial evaluated at the point  $(2, 1)$  counts the number of forests of the graph. We have shown rapid mixing at all fixed points  $(2, 1 + \delta)$ , for  $\delta > 0$ , with a mixing time that depends on  $\delta$ . It would be interesting to consider whether the chains associated with these boundary points mix rapidly for graphs of bounded tree-width.

The structure of this paper is as follows. In the next section, we give the definitions that are necessary for a detailed description of the main theorem. We give the main theorem in Section 3 and then indicate some of its consequences. We present the proofs in Section 4. In Section 5, we extend our results to Glauber dynamics on vertex subsets, that is, on induced subgraphs.

## 2 Definitions

### 2.1 $\lambda$ -multiplicative weight functions

In this subsection, we describe the condition we require on our graph functions  $\mathcal{P}$ . This condition prescribes that the weight function is multiplicative with respect to the operation of disjoint graph union as well as “nearly multiplicative” with respect to the operation of composition via small vertex cuts.

We use the notation  $\hat{\lambda} := \max\{\lambda, 1/\lambda\}$ . For a graph  $G = (V, E)$ , a vertex cut  $K$  is said to separate sets  $V_1$  and  $V_2$  if  $(V_1, K, V_2)$  is a partition of  $V$  and there is no edge of  $E$  that is incident to both a vertex of  $V_1$  and a vertex of  $V_2$ . A partition  $(E_1, E_2)$  of  $E$  is appropriate (for  $K$ ) if  $E_1$  has no edge adjacent to a vertex in  $V_2$  and  $E_2$  has no edge adjacent to a vertex in  $V_1$ .

For fixed  $\lambda > 0$ , we say that the weight function  $w$  is  $\lambda$ -multiplicative, if for any  $G = (V, E)$ , any vertex cut  $K$  that separates sets  $V_1$  and  $V_2$ , and any appropriate partition  $(E_1, E_2)$ , we have

$$\hat{\lambda}^{-|K|} \leq \frac{w((V_1 \cup K, E_1))w((V_2 \cup K, E_2))}{w(G)} \leq \hat{\lambda}^{|K|}. \quad (6)$$

As mentioned above, if  $w$  is  $\lambda$ -multiplicative, then it follows that  $w$  is multiplicative with respect to disjoint union (by taking  $K = \emptyset$ ); furthermore, taking  $V_2 = \emptyset$  implies that the addition or deletion of a few edges in the graph does not change  $w$  wildly.

### 2.2 Examples of valid polynomials

In this subsection, we emphasise specific examples of edge subset weightings and justify that their weight functions are  $\lambda$ -multiplicative.

Let  $G = (V, E)$  be any graph,  $K$  be any vertex cut that separates vertex subsets  $V_1$  and  $V_2$ , and  $(E_1, E_2)$  be any appropriate partition. We define  $G'$  to be the disjoint union of graphs  $(V_1 \cup K, E_1)$  and  $(V_2 \cup K, E_2)$ . We could imagine forming  $G'$  from  $G$  by splitting each vertex in  $K$ , taking incident edges in  $E_1$  with one copy of the vertex and those in  $E_2$  with the other. It is trivial to verify multiplicativity with respect to disjoint union for each of the weight

functions considered below. Therefore, to establish  $\lambda$ -multiplicativity for these weight functions, it will suffice to verify that  $\hat{\lambda}^{-|K|} \leq w(G')/w(G) \leq \hat{\lambda}^{|K|}$ .

First, we observe that the partition function of the *random cluster model* for  $q, \mu > 0$  satisfies the condition. Recalling (2), the relevant weight function is  $w((V, S)) := q^{\kappa(S)} \mu^{|S|}$ . To handle the  $\mu^{|S|}$  factor, note that the graphs  $G$  and  $G'$  have the same number of edges. For the  $q^{\kappa(S)}$  factor, the number of components in  $G'$  can be at most  $\kappa(G) + |K|$  since  $G'$  can be obtained by splitting  $|K|$  vertices of  $G$ . Thus,  $w$  is  $\lambda$ -multiplicative if we take  $\lambda := q$ .

This can also be seen in the context of the *Tutte polynomial* when  $x, y > 1$ . Recalling (3), the relevant weight function is  $w((V, S)) := (x-1)^{r(E)-r(S)}(y-1)^{|S|-r(S)}$ . As before, it is easy to take care of the  $(x-1)^{r(E)-r(S)}$  factor. For the remaining  $((x-1)(y-1))^{-r(S)}$  factor, it is enough to observe that the incidence matrix of  $G$  may be obtained from the incidence matrix of  $G'$  as follows. The matrix for  $G'$  has two rows for each of the vertices in  $K$ , one from  $(V_1 \cup K, E_1)$  and one from  $(V_2 \cup K, E_2)$ . If we replace one of these two rows with the sum of the two rows, we do not alter the rank; if we then delete the other of the two rows, we change the rank by at most 1. Doing this for each vertex in  $K$ , we obtain the incidence matrix for  $G$ , at a total change in the rank  $r$  of the incidence matrix of at most  $|K|$ . Thus,  $w$  is  $\lambda$ -multiplicative if we take  $\lambda := (x-1)(y-1)$ .

Next, we see that the *adjacency-rank polynomial* of Ge and Štefankovič [22] satisfies the condition if  $q, \mu > 0$ . Recalling (4), the relevant weight function is  $w((V, S)) := q^{\text{rk}_2(S)} \mu^{|S|}$ . As before, it is simple to handle the  $\mu^{|S|}$  factor. For the  $q^{\text{rk}_2(S)}$  factor, we note that the adjacency matrix of  $G$  may be formed from the adjacency matrix of  $G'$  by  $|K|$  row additions, followed by  $|K|$  column additions and finally the deletion of  $|K|$  rows and  $|K|$  columns. Since we must delete both rows and columns, the rank  $\text{rk}_2$  of the adjacency matrix may change by up to  $2|K|$ . Thus, in this case,  $w$  is  $\lambda$ -multiplicative when taking  $\lambda := q^2$ .

Now, consider the *multivariate Tutte polynomial* as formulated by Sokal [54], defined for any graph  $G = (V, E)$  and parameters  $q, \vec{v} = \{v_e\}_{e \in E}$  by

$$Z_{\text{Tutte}}(G; q, \vec{v}) := \sum_{S \subseteq E} q^{\kappa(S)} \prod_{e \in S} v_e. \quad (7)$$

Under this expansion,  $w := q^{\kappa(S)} \prod_{e \in S} v_e$  is an edge subset weight function if  $q > 0$  and  $v_e > 0$  for any  $e \in E$  are fixed. We can handle the  $q^{\kappa(S)}$  factor as we did for the random cluster model partition function. For the  $\prod_{e \in S} v_e$  factor, observe that  $G$  and  $G'$  have the same set of edges. Thus,  $w$  is  $\lambda$ -multiplicative when taking  $\lambda := q$ .

Last, we discuss the *U-polynomial* of Noble and Welsh [49], defined for any graph  $G = (V, E)$  and parameters  $y, \vec{x} = \{x_i\}_{i=1}^{|V|}$  by

$$U(G; \vec{x}, y) := \sum_{S \subseteq E} (y-1)^{|S|-r(S)} \prod_{i=1}^{|V|} x_i^{\kappa(i, S)}, \quad (8)$$

where  $\kappa(i, S)$  denotes the number of components of order  $i$  in  $(V, S)$ . If  $y > 1$  and  $x_i > 0$  for all  $i$ , then  $w((V, S)) := (y-1)^{|S|-r(S)} \prod_{i=1}^{|V|} x_i^{\kappa(i, S)}$  gives an edge subset weighting. The  $(y-1)^{|S|-r(S)}$  factor can be handled as above. For the  $\prod_{i=1}^{|V|} x_i^{\kappa(i, S)}$  factor, observe that  $\sum_i |\kappa(i, G) - \kappa(i, G')|$  is at most  $3|K|$ ,

since, if we obtain  $G'$  by splitting the vertices in  $K$ , each time we split a vertex we either change the size of a single component or split a single component into two smaller components. Thus, taking  $x' := \max_i \max\{x_i, x_i^{-1}\}$  and  $y' := \max\{y - 1, (y - 1)^{-1}\}$ , we see that  $w$  is  $\lambda$ -multiplicative when taking  $\lambda := y'x'^3$ .

### 2.3 Glauber dynamics for edge subsets

In this subsection, we define the Markov chain associated with the edge subset expansion formula for  $\mathcal{P}$ . From the formulation in (1), the *single bond flip chain*  $\mathcal{M}$  on a given graph  $G = (V, E)$  is defined as follows. We start with an arbitrary subset  $X_0 \subseteq E$  and repeatedly generate  $X_{t+1}$  from  $X_t$  by running the following experiment.

1. Pick an edge  $e \in E$  uniformly at random and let  $S = X_t \oplus \{e\}$ .
2. Set  $X_{t+1} = S$  with probability  $\frac{1}{2} \min\{1, w((V, S))/w((V, X_t))\}$  and with the remaining probability set  $X_{t+1} = X_t$ .

By convention, we denote the state space of  $\mathcal{M}$  by  $\Omega$  (i.e.  $\Omega = 2^E$ ) and its transition probability matrix by  $P$ . With standard arguments, it can be shown that  $\mathcal{M}$  is a reversible Markov chain that has a unique stationary distribution  $\pi$  satisfying  $\pi(S) \propto w((V, S))$ . Hence, we may use  $\mathcal{M}$  as a Markov chain in MCMC sampling for the following problem.

**PWE( $\mathcal{P}$ ):  $\mathcal{P}$ -WEIGHTED EDGE SUBSETS**

Input: a graph  $G = (V, E)$ .

Output: a subset  $S \subseteq E$  with probability  $w((V, S))/\mathcal{P}(G)$ .

The term *rapidly mixing* applies to a Markov chain that quickly converges to its stationary distribution. We make this precise here. The *total variation distance*  $\|\nu - \nu'\|_{TV}$  between two probability distributions  $\nu$  and  $\nu'$  is defined by  $\|\nu - \nu'\|_{TV} = \frac{1}{2} \sum_{H \in \Omega} |\nu(H) - \nu'(H)|$ . For  $\varepsilon > 0$ , the *mixing time* of a Markov chain  $\mathcal{M}$  (with state space  $\Omega$ , transition matrix  $P$  and stationary distribution  $\pi$ ) is defined as

$$\tau(\varepsilon) := \max_{H \in \Omega} \{\min\{t \mid \|P^t(H, \cdot) - \pi(\cdot)\|_{TV} \leq \varepsilon\}\}.$$

In this paper, we shall say that a chain  $\mathcal{M}$  *mixes rapidly* if, for any fixed  $\varepsilon$ ,  $\tau(\varepsilon)$  is (upper) bounded by a polynomial in the number of vertices of the input graph. This definition for rapid mixing is the one more commonly used in theoretical computer science, whereas often in statistical physics or discrete probability a stricter  $O(n \log n)$  bound is mandated.

## 3 Results

We are now prepared to state the main theorem.

**Theorem 1.** *Let  $G = (V, E)$  where  $|V| = n$ . If  $w$  is  $\lambda$ -multiplicative for some  $\lambda > 0$ , then the mixing time of  $\mathcal{M}$  on  $G$  satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\text{tw}(G)+1)|\log \lambda|} \log(1/\varepsilon)\right)$$

(where  $\text{tw}(G)$  denotes the tree-width of  $G$ ).

In Subsection 2.2, we noted some examples of polynomials that have  $\lambda$ -multiplicative weight functions; thus, Theorem 1 implies that each of their associated Glauber dynamics on edge subsets is rapidly mixing upon graphs of bounded tree-width.

**Corollary 2.** *Let  $G = (V, E)$  where  $|V| = n$ . In the following list, we state conditions on the parameters which guarantee rapid mixing of the single bond flip chain on  $G$  associated with the stated polynomial and weighting. We also state the mixing time bound.*

1. *For fixed  $q, \mu > 0$  and the weighting (2) of  $Z_{RC}(G; q, \mu)$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\text{tw}(G)+1)|\log q|} \log(1/\varepsilon)\right).$$

*Equivalently, for fixed  $x, y > 1$  and the weighting (3) of  $T(G; x, y)$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\text{tw}(G)+1)|\log((x-1)(y-1))|} \log(1/\varepsilon)\right).$$

2. *For fixed  $q, \mu > 0$  and the weighting (4) of  $R_2(G; q, \mu)$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{4+8(\text{tw}(G)+1)|\log q|} \log(1/\varepsilon)\right).$$

3. *For fixed  $q > 0$  and  $v_e > 0$  for all  $e$  and the weighting (7) of  $Z(G; q, \vec{v})$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\text{tw}(G)+1)|\log q|} \log(1/\varepsilon)\right).$$

4. *For fixed  $y > 1$  and  $x_i > 0$  for all  $i$  and the weighting (8) of  $U(G; \vec{x}, \mu)$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\text{tw}(G)+1)|\log(y'x'^3)|} \log(1/\varepsilon)\right)$$

*where  $x' = \max_i \max\{x_i, x_i^{-1}\}$  and  $y' = \max\{y - 1, (y - 1)^{-1}\}$ .*

Here, we remark that Ge and Štefankovič obtained part 1 above and showed part 2 above in the special case of trees. Parts 2–4 directly extend these findings, and our main theorem considerably broadens the scope of mixing time bounds for subset Glauber dynamics on graphs of bounded tree-width.

## 4 Proofs

Let us first give an outline of the proof.

Although our main result is stated in terms of tree-width, we do not treat tree-width directly but instead use linear-width, a more restrictive width parameter introduced by Thomas [56]. This strategy was also employed by Ge and Štefankovič in the two specific cases mentioned above. For any graph



$G = (V, E)$ , an ordering  $(e_1, \dots, e_m)$  of  $E$  has linear-width at most  $\ell$ , if, for each  $i \in \{1, \dots, m\}$ , there are at most  $\ell$  vertices that are incident to both an edge in  $\{e_1, \dots, e_{i-1}\}$  and an edge in  $\{e_i, \dots, e_m\}$ . The *linear-width*  $\text{lw}(G)$  of  $G = (V, E)$  is the smallest integer  $\ell$  such that there is an ordering of  $E$  with linear-width at most  $\ell$ . The motive for using linear-width is that it implies an ordering of the edges which we can then use to define canonical paths between pairs of edge subsets. Then we show that  $\lambda$ -multiplicativity is the general condition under which we can bound the congestion of these canonical paths. The use of canonical paths is a standard technique for obtaining a bound on the mixing time of MCMC methods — see the lecture notes of Jerrum [35] for an expository account of this approach.

The key property we require that relates the linear-width of  $G$  to the more well-studied parameters path-width  $\text{pw}(G)$  and tree-width  $\text{tw}(G)$  of  $G$  is the following set of inequalities, details of which can be found in Bodlaender [11], Chung and Seymour [14], Fomin and Thilikos [21], Ge and Štefankovič [22], and Korach and Solel [42]. For any graph  $G$  on  $n$  vertices,

$$\text{pw}(G) \leq \text{lw}(G) \leq \text{pw}(G) + 1 \leq (\text{tw}(G) + 1)(\lfloor \log_2 n \rfloor + 1) + 1. \quad (9)$$

We follow a canonical paths strategy to bound the mixing time of  $\mathcal{M}$ . Given  $G = (V, E)$ , let  $\sigma = (e_1, \dots, e_m)$  be an ordering of  $E$ . Given  $I, F \in \Omega$ , let  $I \oplus F$  denote the symmetric difference of  $I$  and  $F$ , let  $\sigma[I \oplus F] := (e_{i_1}, \dots, e_{i_k})$  denote the restriction of  $\sigma$  to  $I \oplus F$  (that is,  $\{e_{i_1}, \dots, e_{i_k}\} = I \oplus F$  and  $i_1 < \dots < i_k$ ), and let  $\gamma_{\sigma, I \rightarrow F}$  denote the *canonical path* from  $I$  to  $F$ , defined as

$$\gamma_{\sigma, I \rightarrow F} := (H_0, \dots, H_k),$$

where  $H_0 = I$ ,  $H_j = H_{j-1} \oplus \{e_{i_j}\}$  for all  $j \in \{1, \dots, k\}$  (and hence  $H_k = F$ ). Let  $\Gamma_\sigma = \{\gamma_{\sigma, I \rightarrow F} \mid I, F \in \Omega\}$ .

To bound the mixing time of  $\mathcal{M}$ , we will, for some appropriately chosen  $\sigma$ , bound the *congestion*  $\varrho(\Gamma_\sigma)$  of the canonical paths, which is defined by

$$\varrho(\Gamma_\sigma) := \max_{\substack{(H, H') : \\ P(H, H') > 0}} \left\{ \frac{1}{\pi(H)P(H, H')} \sum_{\substack{I, F : \\ (H, H') \in \gamma_{\sigma, I \rightarrow F}}} \pi(I)\pi(F) |\gamma_{\sigma, I \rightarrow F}| \right\}, \quad (10)$$

where  $|\gamma_{\sigma, I \rightarrow F}|$  denotes the length of the path  $\gamma_{\sigma, I \rightarrow F}$ . The mixing time can then be bounded using the following inequality of Sinclair [53] — see also Diaconis and Stroock [17]: for any set  $\Gamma$  of canonical paths,

$$\tau(\varepsilon) \leq \max_{H \in \Omega} \left\{ \varrho(\Gamma) \cdot \left( \log \frac{1}{\pi(H)} + \log \frac{1}{\varepsilon} \right) \right\}. \quad (11)$$

The remainder of the section is devoted to showing the following.

**Theorem 3.** *Suppose  $G = (V, E)$  has linear-width  $\ell$  and let  $\sigma = (e_1, \dots, e_m)$  be an ordering of  $E$  with linear-width at most  $\ell$ . If  $w$  is  $\lambda$ -multiplicative for some  $\lambda > 0$ , then  $\varrho(\Gamma_\sigma) \leq 2m^2 \hat{\lambda}^{4\ell}$ .*

Theorem 3 immediately implies a good mixing time bound for the Markov chain  $\mathcal{M}$  and hence Theorem 1 follows.

**Corollary 4.** Let  $G = (V, E)$  where  $|E| = m$ . If  $w$  is  $\lambda$ -multiplicative for some  $\lambda > 0$ , then the mixing time of  $\mathcal{M}$  on  $G$  satisfies

$$\tau(\varepsilon) = O\left(m^2 \hat{\lambda}^{4 \text{lw}(G)} \log(1/\varepsilon)\right).$$

*Proof.* Substitute the congestion bound of Theorem 3 into the inequality (11).  $\square$

*Proof of Theorem 1.* Substitute the upper bound on  $\text{lw}(G)$  in (9) into Corollary 4.  $\square$

In the proof of Theorem 3, we will need the following lemma.

**Lemma 5.** Suppose  $G = (V, E)$  has linear-width  $\ell$  and let  $\sigma = (e_1, \dots, e_m)$  be an ordering of  $E$  with linear-width at most  $\ell$ . Suppose  $I, F \in \Omega$  and  $H$  is on  $\gamma_{\sigma, I \rightarrow F}$ . If  $w$  is  $\lambda$ -multiplicative for some  $\lambda > 0$ , then

$$\frac{w((V, I))w((V, F))}{w((V, H))w((V, C))} \leq \hat{\lambda}^{4\ell},$$

where  $C = I \oplus F \oplus H$ .

*Proof.* Since  $H$  is on  $\gamma_{\sigma, I \rightarrow F}$ , we may assume that  $H = H_j$  for some  $j \in \{0, \dots, k\}$ . Let  $Q = \{e_1, \dots, e_{i_j}\}$  and  $\overline{Q} = E \setminus Q$ . Then

$$H = (F \cap Q) \cup (I \cap \overline{Q}) \quad \text{and} \quad C = (I \cap Q) \cup (F \cap \overline{Q}). \quad (12)$$

We can partition  $V$  into three sets as follows. Let  $V_1$  denote the set of vertices that are incident only to edges in  $Q$ ; let  $V_2$  denote the set of vertices that are incident only to edges in  $\overline{Q}$ ; let  $K$  denote the set of remaining vertices, that is, the set of vertices incident to edges in  $Q$  and  $\overline{Q}$ . Note that  $|K|$  is at most the linear-width  $\ell$ .

No vertex  $v_1$  of  $V_1$  is adjacent to a vertex  $v_2$  of  $V_2$ , as otherwise the edge between them would simultaneously be in  $Q$  and  $\overline{Q}$ . This implies that  $K$  is a vertex cut separating  $V_1$  and  $V_2$  with respect to  $G$ , and also with respect to the graphs  $(V, I)$ ,  $(V, F)$ ,  $(V, H)$ ,  $(V, C)$ . Furthermore,  $(I \cap Q, I \cap \overline{Q})$ ,  $(F \cap Q, F \cap \overline{Q})$ ,  $(H \cap Q, H \cap \overline{Q})$ ,  $(C \cap Q, C \cap \overline{Q})$  are edge partitions that are appropriate for  $K$ . Therefore, by the fact that  $w$  is  $\lambda$ -multiplicative and  $|K| \leq \ell$ ,

$$\hat{\lambda}^{-\ell} \leq \frac{w((V_1 \cup K, J \cap Q))w((V_2 \cup K, J \cap \overline{Q}))}{w((V, J))} \leq \hat{\lambda}^{\ell} \quad \text{for } J \in \{I, F, H, C\}.$$

By (12), it follows that

$$\begin{aligned} (V_1 \cup K, H \cap Q) &= (V_1 \cup K, F \cap Q), \\ (V_2 \cup K, H \cap \overline{Q}) &= (V_2 \cup K, I \cap \overline{Q}), \\ (V_1 \cup K, C \cap Q) &= (V_1 \cup K, I \cap Q) \quad \text{and} \\ (V_2 \cup K, C \cap \overline{Q}) &= (V_2 \cup K, F \cap \overline{Q}). \end{aligned}$$

Now, letting  $r$  be

$$w((V_1 \cup K, I \cap Q))w((V_2 \cup K, I \cap \overline{Q}))w((V_1 \cup K, F \cap Q))w((V_2 \cup K, F \cap \overline{Q}))$$

we obtain

$$\frac{w((V, I))w((V, F))}{r} \leq \hat{\lambda}^{2\ell} \quad \text{and} \quad \frac{r}{w((V, H))w((V, C))} \leq \hat{\lambda}^{2\ell},$$

whereby the lemma easily follows.  $\square$

*Proof of Theorem 3.* Let  $(H, H') \in \Omega \times \Omega$  such that  $P(H, H') > 0$ . We will bound the expression within the max of the definition for  $\varrho(\Gamma_\sigma)$ . We let  $\hat{H} = H$  if  $\pi(H) \leq \pi(H')$  and  $\hat{H} = H'$  otherwise. Denote by  $\text{cp}(H, H')$  the set of pairs  $(I, F)$  such that  $(H, H') \in \gamma_{\sigma, I \rightarrow F}$ . We define the function  $\text{inj} : \text{cp}(H, H') \rightarrow \Omega$  by  $(I, F) \rightarrow I \oplus F \oplus \hat{H}$ . Observe that  $\text{inj}$  is an injection, for, given  $J \in \Omega$ , we can determine the unique  $(I, F)$  such that  $\text{inj}(I, F) = J$  by first computing  $J \oplus \hat{H} = I \oplus F$  and then using the ordering  $\sigma$  to recover  $I$  and  $F$ . Since  $w$  is  $\lambda$ -multiplicative, we have by Lemma 5 that

$$\frac{w((V, I))w((V, F))}{w((V, \hat{H}))w((V, \text{inj}(I, F)))} \leq \hat{\lambda}^{4\ell}. \quad (13)$$

Regardless of whether  $\pi(H) \leq \pi(H')$  or  $\pi(H) > \pi(H')$ , a brief calculation yields that  $\pi(H)P(H, H') = \pi(\hat{H})/(2m)$ ; thus,

$$\begin{aligned} \frac{1}{\pi(H)P(H, H')} \sum_{(I, F) \in \text{cp}(H, H')} \pi(I)\pi(F)|\gamma_{\sigma, I \rightarrow F}| \\ = \frac{2m}{\pi(\hat{H})} \sum_{(I, F) \in \text{cp}(H, H')} \pi(I)\pi(F)|\gamma_{\sigma, I \rightarrow F}| \\ \leq \frac{2m^2}{\mathcal{P}(G)} \sum_{(I, F) \in \text{cp}(H, H')} \frac{w((V, I))w((V, F))}{w((V, \hat{H}))} \end{aligned} \quad (14)$$

$$\leq \frac{2m^2}{\mathcal{P}(G)} \sum_{(I, F) \in \text{cp}(H, H')} w((V, \text{inj}(I, F))) \hat{\lambda}^{4\ell} \quad (15)$$

$$\leq 2m^2 \hat{\lambda}^{4\ell}, \quad (16)$$

where (14) follows from the facts  $|\gamma_{\sigma, I \rightarrow F}| \leq m$  and  $\pi(S) = w((V, S))/\mathcal{P}(G)$ , (15) follows from (13), and (16) follows from the fact that  $\text{inj}$  is an injection. Then, substituting the bound (16) into (10), we obtain  $\varrho(\Gamma_\sigma) \leq 2m^2 \hat{\lambda}^{4\ell}$ , as claimed.  $\square$

## 5 Vertex subset Glauber dynamics for bounded tree-width

Until now, we had been considering edge subsets (subgraphs) and Glauber transitions which change one edge at a time. In this section, we modify our methods to treat vertex subsets (induced subgraphs) and transitions that involve one vertex at a time — each such transition can affect many edges, up to the maximum degree of  $G$ . We sketch how to obtain rapid mixing for this process upon graphs of bounded tree-width still with only a modest condition on the base graph polynomials.

A *vertex subset expansion formula* for  $\mathcal{P}$  is written as follows: for any simple graph  $G = (V, E)$ ,

$$\mathcal{P}(G) = \sum_{S \subseteq V} w(G[S]) \quad (17)$$

for some graph function  $w$ , where  $G[S]$  denotes the subgraph of  $G$  induced by  $S$ . If the function  $w$  is non-negative, we refer to (17) as a *vertex subset weighting* for  $\mathcal{P}$  and to  $w$  as its *weight function*. Again, for our results to hold, aside from some other constraints, we need the weight function to be positive on all induced subgraphs.

From the formulation in (17), we define the *single site flip chain*  $\mathcal{M}'$  on a given graph  $G = (V, E)$  as follows. We start with an arbitrary subset  $X_0 \subseteq V$  and repeatedly generate  $X_{t+1}$  from  $X_t$  by running the following experiment.

1. Pick a vertex  $v \in V$  uniformly at random and let  $S = X_t \oplus \{v\}$ .
2. Set  $X_{t+1} = S$  with probability  $\frac{1}{2} \min \{1, w(G[S])/w(G[X_t])\}$  and with the remaining probability set  $X_{t+1} = X_t$ .

We denote the state space of  $\mathcal{M}'$  by  $\Omega'$  (i.e.  $\Omega' = 2^V$ ) and its transition probability matrix by  $P'$ . It can be shown that  $\mathcal{M}'$  is a reversible Markov chain that has a unique stationary distribution  $\pi'$  satisfying  $\pi'(S) \propto w(G[S])$ . Hence, we may use  $\mathcal{M}'$  as a Markov chain in MCMC sampling for the following problem.

**PWV( $\mathcal{P}$ ):  $\mathcal{P}$ -WEIGHTED VERTEX SUBSETS**

Input: a graph  $G = (V, E)$ .

Output: a subset  $S \subseteq V$  with probability  $w(G[S])/\mathcal{P}(G)$ .

We now describe the condition required of the weight function  $w$  in (17). For fixed  $\lambda > 0$ , we say that the weight function  $w$  is *vertex  $\lambda$ -multiplicative*, if for any  $G = (V, E)$  and  $K$  a vertex cut that separates sets  $V_1$  and  $V_2$  with respect to  $G$ , we have

$$\hat{\lambda}^{-|K|} \leq \frac{w(G[V_1])w(G[V_2 \cup K])}{w(G)} \leq \hat{\lambda}^{|K|}. \quad (18)$$

Note that, if  $w$  is vertex  $\lambda$ -multiplicative, then it follows that  $w$  is multiplicative with respect to disjoint union by taking  $K = \emptyset$ ; furthermore, taking  $V_2 = \emptyset$  gives that the addition of a few vertices does not change  $w$  wildly.

The main result of this section is the following.

**Theorem 6.** *Let  $G = (V, E)$  where  $|V| = n$ . If  $w$  is vertex  $\lambda$ -multiplicative for some  $\lambda > 0$ , then the mixing time of  $\mathcal{M}'$  on  $G$  satisfies*

$$\tau(\varepsilon) = O\left(n^{2+4(\text{tw}(G)+1)|\log \lambda|} \log(1/\varepsilon)\right).$$

## 5.1 A sketch of the proof

As before, we do not treat tree-width directly, but instead work with a different width parameter. For any graph  $G = (V, E)$ , an ordering  $(v_1, \dots, v_n)$  of  $V$  has vertex-separation at most  $\ell$ , if, for each  $i \in \{1, \dots, n\}$ , there are at most  $\ell$  vertices in  $\{v_1, \dots, v_{i-1}\}$  that are adjacent to a vertex in  $\{v_i, \dots, v_n\}$ . The

*vertex-separation*  $\text{vs}(G)$  of  $G = (V, E)$  is the smallest integer  $\ell$  such that there is an ordering of  $V$  with vertex-separation at most  $\ell$ . It was shown by Kinnery [41] that the vertex-separation of  $G$  satisfies  $\text{vs}(G) = \text{pw}(G)$ , and so the inequalities in (9) remain relevant.

To bound the mixing time of  $\mathcal{M}'$ , we again follow a canonical paths argument. Given  $G = (V, E)$ , let  $\sigma = (v_1, \dots, v_n)$  be an ordering of  $V$ . Given  $I, F \in \Omega'$ , let  $I \oplus F$  denote the symmetric difference of  $I$  and  $F$ , let  $\sigma[I \oplus F] := (v_{i_1}, \dots, v_{i_k})$  denote the restriction of  $\sigma$  to  $I \oplus F$  (that is,  $\{v_{i_1}, \dots, v_{i_k}\} = I \oplus F$  and  $i_1 < \dots < i_k$ ), and let  $\gamma_{\sigma, I \rightarrow F}$  denote the *canonical path* from  $I$  to  $F$ , defined as

$$\gamma_{\sigma, I \rightarrow F} := (H_0, \dots, H_k),$$

where  $H_0 = I$ ,  $H_j = H_{j-1} \oplus \{v_{i_j}\}$  for all  $j \in \{1, \dots, k\}$  (and hence  $H_k = F$ ). Let  $\Gamma_\sigma = \{\gamma_{\sigma, I \rightarrow F} \mid I, F \in \Omega\}$ . Using inequality (11), our bound on the mixing time again follows from a bound on the *congestion*  $\varrho(\Gamma_\sigma)$ , which is defined analogously to (10).

**Theorem 7.** *Suppose  $G = (V, E)$  has vertex-separation  $\ell$ . Let  $\sigma = (v_1, \dots, v_n)$  be an ordering of  $V$  with vertex-separation at most  $\ell$ . If, for some  $\lambda > 0$ ,  $w$  is vertex  $\lambda$ -multiplicative, then  $\varrho(\Gamma_\sigma) \leq 2n^2 \hat{\lambda}^{4\ell}$ .*

Theorem 7 immediately implies a good mixing time bound for the Markov chain  $\mathcal{M}'$  and hence Theorem 6 also.

**Corollary 8.** *Let  $G = (V, E)$  where  $|V| = n$ . If  $w$  is vertex  $\lambda$ -multiplicative for some  $\lambda > 0$ , then the mixing time of  $\mathcal{M}'$  on  $G$  satisfies*

$$\tau(\varepsilon) = O\left(n^2 \hat{\lambda}^{4 \text{vs}(G)} \log(1/\varepsilon)\right).$$

*Proof.* Substitute the congestion bound in (11) into Theorem 7.  $\square$

*Proof of Theorem 6.* Substitute the upper bound on  $\text{vs}(G) = \text{pw}(G)$  in (9) into Corollary 8.  $\square$

We omit the proof of Theorem 7 as it is similar to that of Theorem 3, but give the details for the analogue of Lemma 5.

**Lemma 9.** *Suppose  $G = (V, E)$  has vertex-separation  $\ell$  and let  $\sigma = (v_1, \dots, v_n)$  be an ordering of  $V$  with vertex-separation at most  $\ell$ . Suppose  $I, F \in \Omega'$  and  $H$  is on  $\gamma_{\sigma, I \rightarrow F}$ . If  $w$  is vertex  $\lambda$ -multiplicative for some  $\lambda > 0$ , then*

$$\frac{w(G[I])w(G[F])}{w(G[H])w(G[C])} \leq \hat{\lambda}^{4\ell},$$

where  $C = I \oplus F \oplus H$ .

*Proof.* Since  $H$  is on  $\gamma_{\sigma, I \rightarrow F}$ , we may assume that  $H = H_j$  for some  $j \in \{0, \dots, k\}$ . Let  $Q = \{v_1, \dots, v_{i_j}\}$ , and  $\overline{Q} = V \setminus Q$ . Then

$$H = (F \cap Q) \cup (I \cap \overline{Q}) \quad \text{and} \quad C = (I \cap Q) \cup (F \cap \overline{Q}). \quad (19)$$

We can partition  $V$  into three sets as follows. Let  $V_1$  denote the set of vertices  $\overline{Q}$ ; let  $V_2$  denote the subset of  $Q$  containing vertices adjacent only to other vertices

of  $Q$ ; and let  $K$  denote the set of remaining vertices, that is, the set of vertices of  $Q$  incident to vertices of  $V_1$ . Note that  $|K|$  is at most the vertex-separation  $\ell$ .

Clearly,  $K$  is a vertex cut separating  $V_1$  and  $V_2$  with respect to  $G$  and also with respect to the graphs  $G[I]$ ,  $G[F]$ ,  $G[H]$ ,  $G[C]$ . Therefore, by the fact that  $w$  is vertex  $\lambda$ -multiplicative, and noting that  $V_2 \cup K = Q$ ,

$$\hat{\lambda}^{-\ell} \leq \frac{w(G[Q \cap J])w(G[\overline{Q} \cap J])}{w(G[J])} \leq \hat{\lambda}^\ell \quad \text{for } J \in \{I, F, H, C\}.$$

By (19), it follows that  $H \cap Q = F \cap Q$ ,  $H \cap \overline{Q} = I \cap \overline{Q}$ ,  $C \cap Q = I \cap Q$  and  $C \cap \overline{Q} = F \cap \overline{Q}$ . Now, letting  $r = w(G[Q \cap I])w(G[\overline{Q} \cap I])w(G[Q \cap F])w(G[\overline{Q} \cap F])$ , we obtain that

$$\frac{w(G[I])w(G[F])}{r} \leq \hat{\lambda}^{2\ell} \quad \text{and} \quad \frac{r}{w(G[H])w(G[C])} \leq \hat{\lambda}^{2\ell},$$

whereby the lemma easily follows.  $\square$

## 5.2 An example of a vertex subset chain

Recalling (5), for fixed  $x, y > 1$ ,  $w(G[S]) := (x-1)^{\text{rk}_2(S)}(y-1)^{|V|-\text{rk}_2(S)}$  gives a vertex subset weighting for  $q(G; x, y)$ . With arguments very similar to those given in Subsection 2.2, it is not difficult to verify that this weight function is vertex  $\lambda$ -multiplicative. So, by Theorem 6, it follows that a natural Markov chain derived from the bivariate interlace polynomial — a chain which has not been studied extensively, as far as we are aware — mixes rapidly on tree-width-bounded graphs.

**Corollary 10.** *Let  $G = (V, E)$  where  $|V| = n$ . If  $x, y > 1$  are fixed, then for the single site flip chain on  $G$  associated with the weighting (5) of  $q(G; x, y)$ , the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{2+8(\text{tw}(G)+1)|\log((x-1)/(y-1))|} \log(1/\varepsilon)\right).$$

We believe that it would be of wider interest to study further properties of this single site flip chain on general graphs, in particular to compare it with known results on the random cluster, Potts and Ising models.

## 6 Conclusion

In this work, we have developed a new general framework of graph polynomials and Markov chains defined via subset expansion formulae for these polynomials, and demonstrated that their dynamics mix rapidly for graphs of bounded tree-width. On a graph  $G$  with  $n$  vertices, we have shown a mixing time of order  $n^{O(1)}e^{O(\text{pw}(G))} = n^{O(\text{tw}(G))}$ . Our results apply to many of the most prominent and well-known polynomials in the field. The mixing times of our processes have, respectively, exponential and super-exponential dependencies upon path-width and tree-width. We ask whether this can be improved, in particular, to achieve something akin to fixed-parameter tractability in terms of tree-width.

## References

- [1] M. Aigner and H. van der Holst. Interlace polynomials. *Linear Algebra Appl.*, 377:11–30, 2004.
- [2] N. Alon, A. Frieze, and D. Welsh. Polynomial time randomised approximation schemes for the Tutte polynomial of dense graphs. In *35th Annual Symposium on Foundations of Computer Science (Santa Fe, NM, 1994)*, pages 24–35. IEEE Comput. Soc. Press, Los Alamitos, CA, 1994.
- [3] N. Alon, A. Frieze, and D. Welsh. Polynomial time randomized approximation schemes for Tutte-Gröthendieck invariants: the dense case. *Random Structures Algorithms*, 6(4):459–478, 1995.
- [4] A. Andrzejak. An algorithm for the Tutte polynomials of graphs of bounded treewidth. *Discrete Math.*, 190(1-3):39–54, 1998.
- [5] R. Arratia, B. Bollobás, and G. B. Sorkin. A two-variable interlace polynomial. *Combinatorica*, 24(4):567–584, 2004.
- [6] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields*, 131(3):311–340, 2005.
- [7] G. D. Birkhoff. A determinant formula for the number of ways of coloring a map. *Ann. of Math. (2)*, 14(1-4):42–46, 1912/13.
- [8] M. Bläser and C. Hoffmann. Fast evaluation of interlace polynomials on graphs of bounded treewidth. To appear in *Algorithmica*. doi:10.1007/s00453-010-9439-4.
- [9] M. Bläser and C. Hoffmann. On the complexity of the interlace polynomial. In S. Albers and P. Weil, editors, *25th International Symposium on Theoretical Aspects of Computer Science (STACS 2008)*, volume 1 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97–108, Dagstuhl, Germany, 2008. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [10] M. Bläser and C. Hoffmann. Fast evaluation of interlace polynomials on graphs of bounded treewidth. In A. Fiat and P. Sanders, editors, *ESA*, volume 5757 of *Lecture Notes in Computer Science*, pages 623–634. Springer, 2009.
- [11] H. L. Bodlaender. Some classes of graphs with bounded treewidth. *Bulletin of the EATCS*, 36:116–125, 1988.
- [12] H. L. Bodlaender. Treewidth: characterizations, applications, and computations. In *Graph-theoretic concepts in computer science*, volume 4271 of *Lecture Notes in Comput. Sci.*, pages 1–14. Springer, Berlin, 2006.
- [13] M. Bordewich. Approximating the number of acyclic orientations for a class of sparse graphs. *Combin. Probab. Comput.*, 13(1):1–16, 2004.
- [14] F. R. K. Chung and P. D. Seymour. Graphs with small bandwidth and cutwidth. *Discrete Math.*, 75(1-3):113–119, 1989. *Graph theory and combinatorics* (Cambridge, 1988).

- [15] B. Courcelle. A multivariate interlace polynomial and its computation for graphs of bounded clique-width. *Electron. J. Combin.*, 15(1):Research Paper 69, 36, 2008.
- [16] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.
- [17] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.
- [18] J. Ding, E. Lubetzky, and Y. Peres. Mixing time of critical Ising model on trees is polynomial in the height. *Comm. Math. Phys.*, 295(1):161–207, 2010.
- [19] J. A. Ellis-Monaghan and C. Merino. Graph polynomials and their applications I: The Tutte polynomial. In M. Dehmer, editor, *Structural Analysis of Complex Networks*, pages 219–255. Birkhäuser Boston, 2011.
- [20] J. A. Ellis-Monaghan and C. Merino. Graph polynomials and their applications II: Interrelations and interpretations. In M. Dehmer, editor, *Structural Analysis of Complex Networks*, pages 257–292. Birkhäuser Boston, 2011.
- [21] F. V. Fomin and D. M. Thilikos. A 3-approximation for the pathwidth of Halin graphs. *J. Discrete Algorithms*, 4(4):499–510, 2006.
- [22] Q. Ge and D. Štefankovič. A graph polynomial for independent sets of bipartite graphs. *CoRR*, abs/0911.4732, 2009.
- [23] Q. Ge and D. Štefankovič. A graph polynomial for independent sets of bipartite graphs. In K. Lodaya and M. Mahajan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010)*, volume 8 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 240–250, Dagstuhl, Germany, 2010. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [24] R. J. Glauber. Time-dependent statistics of the Ising model. *J. Mathematical Phys.*, 4:294–307, 1963.
- [25] L. A. Goldberg and M. Jerrum. Personal communication.
- [26] L. A. Goldberg and M. Jerrum. Approximating the partition function of the ferromagnetic potts model. In S. Abramsky, C. Gavoille, C. Kirchner, F. M. auf der Heide, and P. G. Spirakis, editors, *ICALP (1)*, volume 6198 of *Lecture Notes in Computer Science*, pages 396–407. Springer, 2010.
- [27] L. A. Goldberg, M. Jerrum, and M. Karpinski. The mixing time of Glauber dynamics for coloring regular trees. *Random Structures Algorithms*, 36(4):464–476, 2010.
- [28] G. Gordon and L. Traldi. Generalized activities and the Tutte polynomial. *Discrete Math.*, 85(2):167–176, 1990.
- [29] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.



- [30] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [31] P. Hliněný, S.-i. Oum, D. Seese, and G. Gottlob. Width Parameters Beyond Tree-width and their Applications. *The Computer Journal*, 51(3):326–362, 2008.
- [32] M. L. Huber. Simulation reductions for the Ising model. *ArXiv e-prints*, arXiv:0908.2151 [math.PR], 2009.
- [33] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31:253–258, 1925. 10.1007/BF02980577.
- [34] F. Jaeger, D. L. Vertigan, and D. J. A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Math. Proc. Cambridge Philos. Soc.*, 108(1):35–53, 1990.
- [35] M. Jerrum. *Counting, sampling and integrating: algorithms and complexity*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.
- [36] M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for Ising model (extended abstract). In M. Paterson, editor, *ICALP*, volume 443 of *Lecture Notes in Computer Science*, pages 462–475. Springer, 1990.
- [37] M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993.
- [38] V. F. R. Jones. A polynomial invariant for knots via von Neumann algebras. *Bull. Amer. Math. Soc. (N.S.)*, 12(1):103–111, 1985.
- [39] D. R. Karger. A randomized fully polynomial time approximation scheme for the all terminal network reliability problem. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, STOC ’95, pages 11–17, New York, NY, USA, 1995. ACM.
- [40] D. R. Karger. A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem. *SIAM Rev.*, 43(3):499–522, 2001.
- [41] N. G. Kinnersley. The vertex separation number of a graph equals its path-width. *Inform. Process. Lett.*, 42(6):345–350, 1992.
- [42] E. Korach and N. Solel. Tree-width, path-width, and cutwidth. *Discrete Appl. Math.*, 43(1):97–101, 1993.
- [43] J. A. Makowsky. From a zoo to a zoology: towards a general theory of graph polynomials. *Theory Comput. Syst.*, 43(3-4):542–562, 2008.
- [44] J. A. Makowsky and J. P. Mariño. Farrell polynomials on graphs of bounded tree width. *Adv. in Appl. Math.*, 30(1-2):160–176, 2003. Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).
- [45] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures Algorithms*, 31(2):134–172, 2007.

- [46] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [47] S. D. Noble. Evaluating the Tutte polynomial for graphs of bounded tree-width. *Combin. Probab. Comput.*, 7(3):307–321, 1998.
- [48] S. D. Noble. Evaluating a weighted graph polynomial for graphs of bounded tree-width. *Electron. J. Combin.*, 16(1):Research Paper 64, 14, 2009.
- [49] S. D. Noble and D. J. A. Welsh. A weighted graph polynomial from chromatic invariants of knots. *Ann. Inst. Fourier (Grenoble)*, 49(3):1057–1087, 1999. Symposium à la Mémoire de François Jaeger (Grenoble, 1998).
- [50] J. G. Oxley and D. J. A. Welsh. The Tutte polynomial and percolation. In *Graph theory and related topics (Proc. Conf., Univ. Waterloo, Waterloo, Ont., 1977)*, pages 329–339. Academic Press, New York, 1979.
- [51] R. B. Potts. Some generalized order-disorder transformations. *Proc. Cambridge Philos. Soc.*, 48:106–109, 1952.
- [52] D. Randall. Rapidly mixing Markov chains with applications in computer science and physics. *Computing in Science Engineering*, 8(2):30–41, March–April 2006.
- [53] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992.
- [54] A. D. Sokal. The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 173–226. Cambridge Univ. Press, Cambridge, 2005.
- [55] P. Tetali, J. C. Vera, E. Vigoda, and L. Yang. Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. In M. Charikar, editor, *SODA*, pages 1646–1656. SIAM, 2010.
- [56] R. Thomas. Tree-decompositions of graphs, 1996. Lecture notes; available at <http://www.math.gatech.edu/~thomas/tree.ps>.
- [57] W. T. Tutte. A contribution to the theory of chromatic polynomials. *Canadian J. Math.*, 6:80–91, 1954.
- [58] D. J. A. Welsh. *Complexity: knots, colourings and counting*, volume 186 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993.
- [59] D. J. A. Welsh and C. Merino. The Potts model and the Tutte polynomial. *J. Math. Phys.*, 41(3):1127–1152, 2000. Probabilistic techniques in equilibrium and nonequilibrium statistical physics.